

Marta Woźniak

ALGORYTMY GENETYCZNE W AUDYCIE I KONTROLI WEWNĘTRZNEJ

Wprowadzenie

W związku z coraz większym zainteresowaniem tematyką algorytmów genetycznych oraz możliwościami zastosowania ich do rozstrzygania wielu problemów praktycznych nasuwa się pytanie, czy można je wykorzystać do przeprowadzania audytu i kontroli wewnętrznej.

Celem niniejszego opracowania jest przedstawienie – na przykładzie trzech przypadków – zastosowania algorytmów do audytu i kontroli wewnętrznej. Pierwszy z nich dotyczy problemu komiwojażera, drugi – prognozowania popytu, a trzeci – modelu algorytmu genetycznego Drakosa¹, znajdującego zastosowanie w ochronie przed włamaniami do bankowych systemów komputerowych.

Zastosowanie algorytmów genetycznych w odniesieniu do problemu komiwojażera

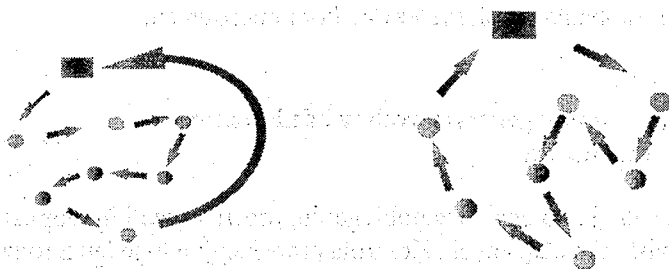
Jak wiadomo, prewencja polega na zapobieganiu, profilaktyce i niedopuszczaniu do powstawania zjawisk niepożądanych. Kontrola prewencyjna to jeden z rodzajów kontroli wewnętrznej. Tradycyjnie kontrola sprawowana w organizacjach (z reguły zwana kontrolą wewnętrzną) dzieli się na kontrolę wewnętrzną instytucjonalną i kontrolę wewnętrzną funkcjonalną. Kontrola instytucjonalna dokonywana jest przez jednostkę kontroli wewnętrznej. Z kolei kontrola funkcjonalna to „kontrola wbudowana w proces biznesowy”. Jest ona sprawowana przez każdego pracownika w zakresie jakości

¹ T. D. Gwiazda, *Algorytmy genetyczne. Zastosowanie w finansach*, Warszawa 1998.

i poprawności wykonywanych przez niego czynności, przez jego bezpośredniego zwierzchnika lub inne wyznaczone osoby czy jednostki (departament, biuro, wydział, zespół lub samodzielne stanowisko) w ramach rutynowych działań. W szczególności może ona być zautomatyzowana i realizowana za pomocą środków technicznych i programowych (tj. sprzętu i oprogramowania). Celem tej kontroli jest zapewnienie zgodności wykonywanych czynności z procedurami, standardami, i innymi uregulowaniami oraz bieżące oddziaływanie i reagowanie na uchybienia².

Nieprawidłowości mogą występować na każdym szczeblu organizacyjnym firmy. Szczególnym przykładem są oszustwa i błędy popełniane przez pracowników, których zadaniem jest proces transportowania przesyłek, przedmiotów czy produktów. Do nieprawidłowości dochodzi zwłaszcza wtedy, gdy proces transportu odbywa się na dużych odległościach i z koniecznością dostarczania określonego towaru do kilku punktów niekorzystnie w stosunku do siebie rozlokowanych. Może tu zaistnieć sytuacja narażająca firmę na straty wynikające z wielu czynników. Jednym z nich jest zawyżony koszt transportu, wynikający z braku kontroli trasy, po której porusza się osoba transportująca. W tym przypadku może dochodzić do umyślnego spowodowania strat finansowych, wynikającego ze świadomego działania pracownika lub też do błędu wynikającego z niezajomości trasy. Innym czynnikiem jest zbyt długi proces dostarczania określonych towarów do odbiorców, co niekiedy wynika z nieprawidłowego ułożenia trasy transportu. Następnym czynnikiem powodującym straty finansowe, również wynikającym z nieprawidłowo ułożonej trasy przejazdu, są opóźnienia w dostarczaniu kolejnych przesyłek spowodowane zbyt długim wykonywaniem wcześniejszego zadania transportowego.

Problem odpowiedniego ułożenia trasy przejazdu nosi nazwę problemu komiwojażera. Polega on na takim dobraniu trasy przejazdu, aby koszt odbytej podróży był jak najmniejszy, a trasa najkrótsza. Wyznaczona trasa powinna się zaczynać i kończyć w jednym miejscu. Ilustrację problemu komiwojażera przedstawia rysunek 1.



Rysunek 1. Problem komiwojażera

Źródło: opracowanie własne.

² K. Czerwiński, H. Grocholski, *Podstawy audytu wewnętrznego*, Szczecin 2003; E. J. Saunders, *Audyt i kontrola wewnętrzna w przedsiębiorstwach*, Częstochowa 2002; Rekomendacja dotycząca kontroli wewnętrznej banku, dokument opublikowany przez Narodowy Bank Polski, Komisję Nadzoru Bankowego, Generalny Inspektorat Nadzoru Bankowego na: http://www.nbp.pl/Publikacje/nadzor_bankowy/pdf/rekomendacja_h.pdf (28.09.2006).

Na rysunku 1 pokazano, że istnieje wiele możliwości ułożenia trasy przejazdu. Punkt startowy i jednocześnie końcowy przedstawiony jest za pomocą prostokąta. Kolejne miasta oznaczone zostały kropkami. Liczba dostępnych możliwości może zostać w każdej chwili obliczona w prosty sposób, z wykorzystaniem elementów kombinatoryki, ponieważ wiąże się to jedynie z obliczeniem liczby możliwości kolejności odwiedzin miast na drodze przejazdu.

Istnieje stosowne oprogramowanie, które układa trasę przejazdu. Do wyznaczenia tras przejazdu stosowane są różne sposoby optymalizacji. Wykorzystywane jest np. obliczanie odległości między poszczególnymi miastami oraz szybkość poruszania się na odpowiednich trasach. Brane są także pod uwagę inne okoliczności, takie jak np. wiadukt na trasie, pod którym nie przejedzie zbyt wysoki pojazd transportujący. W niektórych programach wykorzystywany jest system GPS (*Global Positioning System*). Niekiedy możliwa jest dodatkowa korekta trasy lub wytyczenie nowej w przypadku zaistnienia niekorzystnego zdarzenia losowego, np. wypadku drogowego blokującego miejsce przejazdu i tym samym uniemożliwiającego transport po wcześniej wytyczonej trasie.

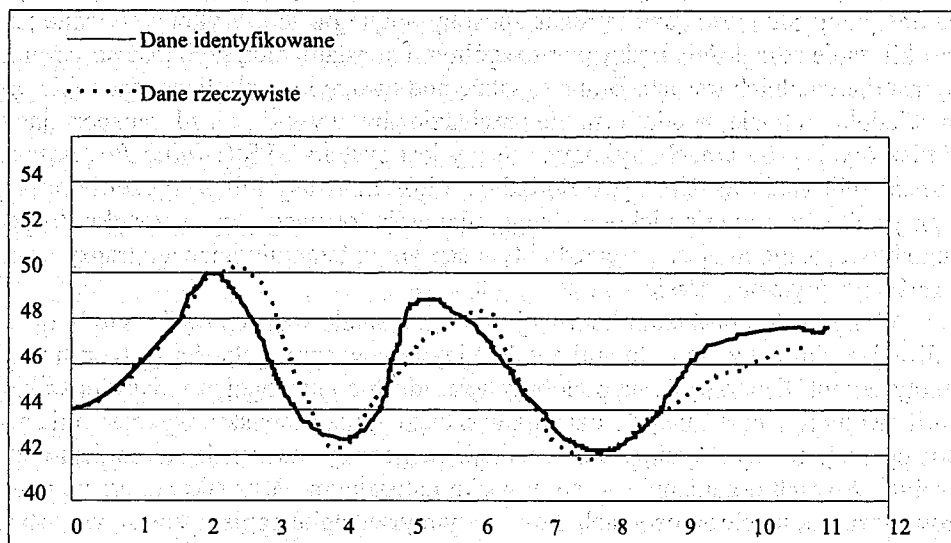
W przypadku problemu komiwojażera można także zastosować metody optymalizacji związane z algorytmami ewolucyjnymi, a w szczególności z algorytmami genetycznymi. Opierają się na podobnych zasadach co mechanizmy darwinowskiej teorii ewolucji, z tą różnicą, że działają w sztucznym środowisku, wykorzystują zasadę operacji genetycznych, proces selekcji, ocenę przystosowania, analogicznie do ewolucji żywych organizmów w środowisku naturalnym. Wszystkie operacje przeprowadzane są na chromosomach, które w tym przypadku zapisywane są w postaci ciągu zer i jedynek. Zastosowanie algorytmu genetycznego pozwala na znalezienie rozwiązania w krótkim czasie, nawet przy dużej liczbie miast.

Zastosowanie algorytmów genetycznych w odniesieniu do prognozowania popytu

Zastosowanie algorytmów genetycznych zostało sprawdzone także w przypadku prognozowania popytu. Według G. Chodaka i W. Kwaśnickiego³, postać funkcji popytu powinna wynikać ze znajomości zarówno wewnętrznego, jak i zewnętrznego środowiska firm. Oznacza to w praktyce, że nie jest możliwe stworzenie uniwersalnej funkcji, którą można by było stosować chociażby w odniesieniu do przedsiębiorstw różnych branż. Przeprowadzone badania oparte były na zależności przewidywanej wartości sprzedaży od popytu, czasu (okresowość sprzedaży) i trendu liniowego. W związku z tym konieczna była identyfikacja dziewięciu parametrów występujących w badanej funkcji. W przypadku tak złożonej funkcji – głównie ze względu na konieczność stosowania skomplikowanych operacji matematycznych – korzystne wydaje się zastosowanie algorytmu genetycznego. Opisane badania zostały przeprowadzone dla rzeczywistych danych pochodzących z dwóch firm. Dla pierwszej firmy, w której sprzedaż była okresowa z rosnącym trendem liniowym, krzywa sprzedaży

³ G. Chodak, W. Kwaśnicki, *Zastosowanie algorytmów genetycznych w prognozowaniu popytu*, http://prev.prawo.uni.wroc.pl/~kwasnicki/todownload/prognozowanie_popytu.pdf (28.09.2006).

wyznaczona z użyciem algorytmu genetycznego niemal pokrywała się z krzywą rzeczywistej sprzedaży. W przypadku drugiej firmy w jednym z kolejnych miesięcy nastąpił gwałtowny wzrost popytu. Prognozowana krzywa sprzedaży różni się w tym miejscu od rzeczywistej krzywej sprzedaży, jednak w pozostałym okresie dobrze ją odwzorowuje.

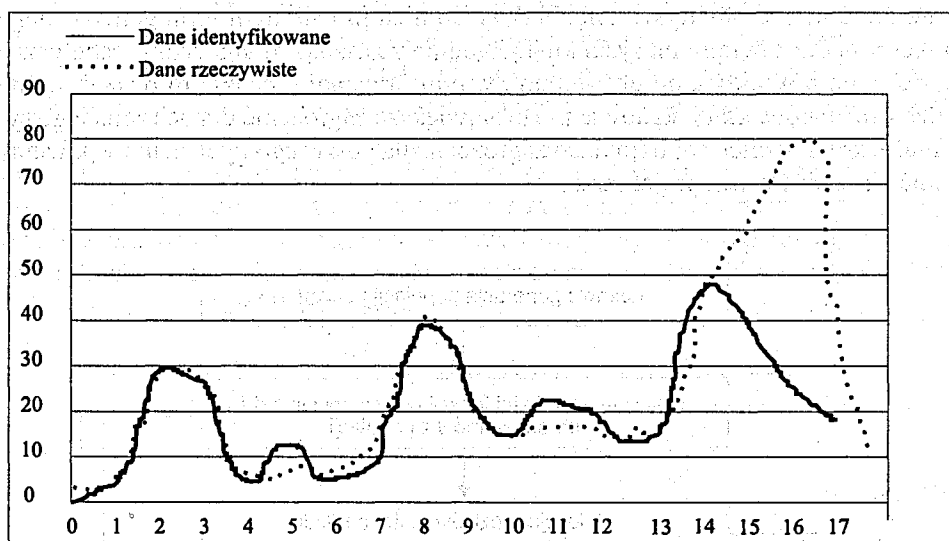


Rysunek 2. Sprzedaż rzeczywista towaru 1 oraz zidentyfikowana za pomocą AG

Źródło: G. Chodak, W. Kwaśnicki, *Zastosowanie algorytmów genetycznych w prognozowaniu popytu*, http://prev.prawo.uni.wroc.pl/~kwasnicki/todownload/prognozowanie_popytu.pdf

Biorąc pod uwagę powyższe dane, można stwierdzić, że stosowanie algorytmów genetycznych w przypadku identyfikacji parametrów funkcji popytu jest uzasadnione. Identyfikacja parametrów funkcji sprzedaży z wykorzystaniem opisanych algorytmów zgodnie z artykułem G. Chodaka i W. Kwaśnickiego daje większe możliwości prognozowania popytu w przypadku zaistnienia sezonowości niż powszechnie znane i często wykorzystywane metody, także jak analiza harmoniczna czy adaptacyjny model Wintersa⁴. Daje to również możliwości wykorzystania tego typu algorytmów w kontroli wewnętrznej zarówno w celu kontrolowania, jak i poprawy prognozowania sprzedaży w przedsiębiorstwie.

⁴ Z. Sarjusz-Wolski, *Ilościowe metody zarządzania logistycznego w przedsiębiorstwie*, Warszawa 1997.



Rysunek 3. Sprzedaż rzeczywista towaru 2 oraz zidentyfikowana za pomocą AG

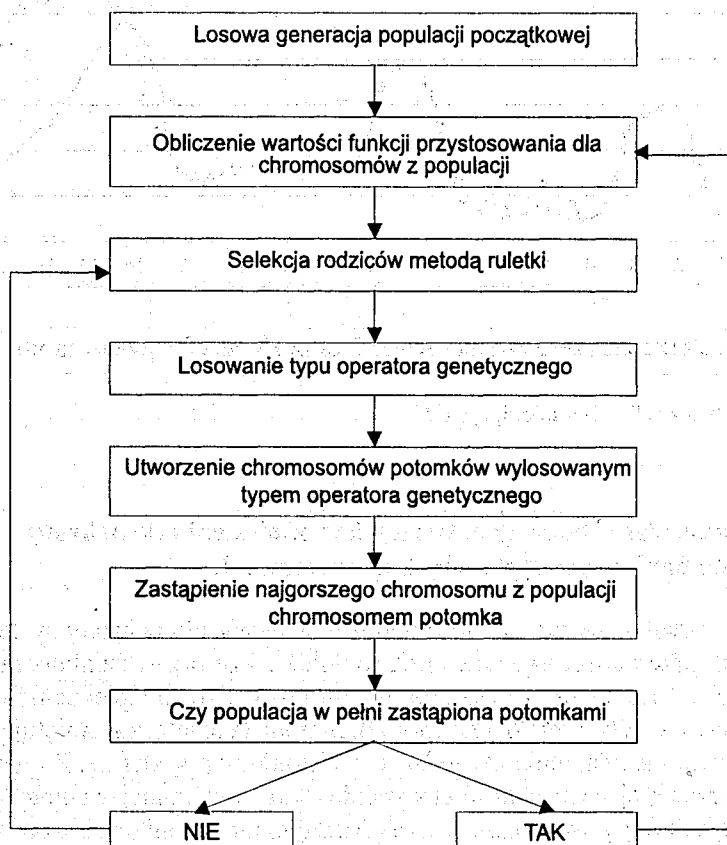
Źródło: G. Chodak, W. Kwaśnicki, *op. cit.*

Zastosowanie algorytmów genetycznych w odniesieniu do ochrony przed włamaniami do bankowych systemów komputerowych

Algorytmy genetyczne mogą być również stosowane dla ochrony systemów komputerowych przed atakami z zewnątrz, również także przed działaniami (błędami) pracowników firm. Jedną z tego typu metod jest nadzór bezpieczeństwa – analiza śladów (ang. *Security Audit Trail Analysis*). Metoda ta działa, zapamiętując wszystkie lub tylko wybrane zdarzenia zachodzące w chronionym systemie. Po tym następuje analiza zdarzeń mająca na celu wychwycenie ataku. Tak działająca metoda nazywana jest metodą ochrony *post factum*. W związku z tym nie chroni ona bezpośrednio przed atakiem, ale pozwala na naprawę ewentualnych szkód przez niego wyrządzonych, a niekiedy nawet przyczynia się do ujęcia przestępcy⁵. Tak działające metody można podzielić na trzy klasy: metody statystyczne lub sieci neuronowe, systemy ekspertowe oraz wnioskowanie modelowe. Metody wnioskowania modelowego polegają na opracowaniu scenariuszy ataku – modeli. W praktyce oznacza to, że standardowe czynności wykonywane w systemie nie stanowią zagrożenia, natomiast jeżeli wystąpią w określonym scenariuszu, są wykrywane jako próba ataku, czyli jako zagrożenie. Problem sprawdzenia, czy w łańcuchu zapamiętanych czynności znajdują się operacje, które razem składają się na scenariusz ataku, należy do klasy problemów NP-zupełnych. W związku z tym do rozstrzygnięcia powyższego problemu można

⁵ T. D. Gwiazda, *op. cit.*

zastosować metody heurystyczne. Należy do nich model algorytmu genetycznego Drakosa, przedstawiony na rysunku 4. Zgodnie z zasadą, że algorytmy genetyczne oparte są na poszukiwaniu ekstremum, w tym przypadku będzie to poszukiwanie podzbioru ataków, który stanowić może największe zagrożenie dla ochranianego systemu. Z kolei różnice w poziomach zagrożenia dla poszczególnych ataków powinny zostać określone przez projektanta.

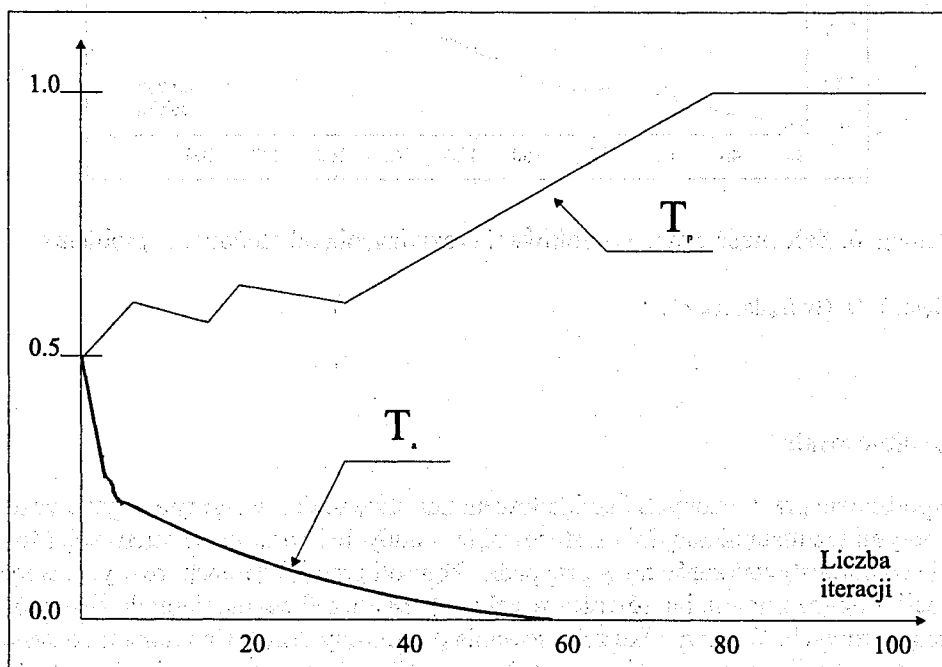


Rysunek 4. Model algorytmu genetycznego Drakosa

Źródło: T. D. Gwiazda, *op. cit.*

Testy dla tej metody przeprowadzono po wprowadzeniu podziału na cztery typy potencjalnych użytkowników systemu: początkujący, zwykły, zawodowy programista oraz superekspert. Badania prowadzone były na rzeczywistych łańcuchach, do których „sztucznie włożono” podzbiory czynności, których zadaniem było symulowanie ataku. Dla oceny wyników wprowadzone zostały dwie wielkości: T_p , oznaczająca

liczbę chromosomów, których geny odpowiadające pojawiającym się atakom były równe 1, oraz T_a , oznaczająca liczbę chromosomów, których geny odpowiadające niezaistniałym atakom były równe 1. Wielkości te wskazują odchylenie uzyskanego wyniku od idealnego, w związku z tym najlepsza sytuacja byłaby wtedy, gdy wartość $T_p = 1$ oraz $T_a = 0$. W praktyce oznaczałoby to, że wszystkie pojawiające się ataki zostały wychwycone i dodatkowo nie było „fałszywego alarmu” spowodowanego rozpoznaniem ataku, który w rzeczywistości nie miał miejsca. Na rysunku 5 pokazane zostały wyniki testów w zależności od liczby iteracji. Rysunek 6 pokazuje natomiast czas potrzebny do znalezienia wyniku w zależności od skomplikowania problemu, czyli od liczby zdefiniowanych scenariuszy.

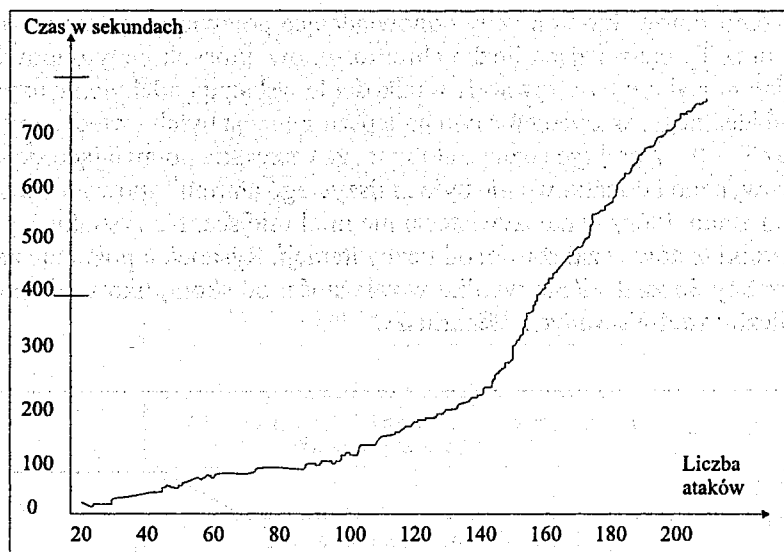


Rysunek 5. Uśrednione wyniki testów

Źródło: T. D. Gwiazda, *op. cit.*

Biorąc pod uwagę wyniki testów⁶, można stwierdzić, że stosowanie algorytmów genetycznych dla ochrony przed atakami systemów komputerowych jest uzasadnione. Metoda ta może być również z powodzeniem wykorzystywana do ochrony przed włamaniami do bankowych systemów komputerowych, a tym samym przed utratą cennych informacji.

⁶ T. D. Gwiazda, *op. cit.*



Rysunek 6. Zależność czasu poszukiwania rozwiązania od złożoności problemu

Źródło: T. D. Gwiazda, *op. cit.*

Podsumowanie

Na podstawie przytoczonych przykładów można stwierdzić, że algorytmy genetyczne z dobrymi rezultatami znajdują zastosowanie w audycie i kontroli wewnętrznej. Można je wykorzystywać zarówno w przypadku kontroli procesów transportowych, weryfikacji prognoz popytu, jak również w celu ochrony przed włamaniami do systemów komputerowych. W przypadku zastosowania do ochrony przed włamaniami do systemów komputerowych, należy pamiętać, że możliwe jest zastosowanie algorytmów genetycznych nie tylko w przypadku banków, ale wszelkich instytucji, w których kradzież informacji może pociągać za sobą konsekwencje, np. finansowe.